Zero123++: a Single Image to Consistent Multi-view Diffusion Base Model

Abstract

We report Zero123++, an image-conditioned diffusion model for generating 3D-consistent multi-view images from a single input view. To take full advantage of pretrained 2D generative priors, we develop various conditioning and training schemes to minimize the effort of finetuning from off-the-shelf image diffusion models such as Stable Diffusion. Zero123++ excels in producing high-quality, consistent multi-view images from a single image, overcoming common issues like texture degradation and geometric misalignment. Furthermore, we showcase the feasibility of training a ControlNet on Zero123++ for enhanced control over the generation process. The code is available at https://github.com/SUDO-AI-3D/zero123plus.